

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## The Evolution and Impact of Library Data

### Conference or Workshop Item

How to cite:

Killick, Selena (2021). The Evolution and Impact of Library Data. In: 14th International Conference on Performance Measurement in Libraries, 2-4 Nov 2021, Online.

For guidance on citations see [FAQs](#).

© 2021 Selena Killick



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](https://oro.open.ac.uk)

# The Evolution and Impact of Library Data

Selena Killick

*The Open University*

## Introduction

As long as there have been libraries, there have been library statistics. It is a powerful tool in our quest to both prove our value and improve our services. Our data collection practices have consequences which can be both positive and negative for our services and our users. Often the consequences of our practices are unknown and unpredictable. This paper provides an overview of approaches to data collection within the profession and considers some of the effects measurement has had on libraries. Trends in library data measurement practices based on published works over the past three decades are presented along with projections on future trends in library data measurement. Real life examples of library data measurement practices are outlined based on the author's experiences within the sector, with specific case studies from The Open University in the United Kingdom. The purpose of this paper is to explore some of the key data practices from the past thirty years and discuss some of their unintended consequences to inform data collection practices of the future.

## Introduction to The Open University

The Open University (OU) is the UK's largest academic institution dedicated to distance learning, with over 210,000 current students across the whole of the UK and Ireland. The University was first established in 1969 with a mission to open up education for all. The OU was purposely designed as a distance learning provider and students study predominantly online. Over 32,000 OU students have a declared disability. The University does not require students to declare a disability, but it does use this information if the student requires any necessary learning adjustments. As such, the OU is the largest provider of higher education for people with disabilities; accessibility of information and learning materials is business critical.

The OU does have a physical library based at the Milton Keynes campus; however customers primarily use online services. 100% of the journal collections and around 80% of the books are electronic. The online library attracts half a million unique users and over 10 million page views per annum. The library helpdesk is available via webchat 24/7/365 where the Library staff provide help and support to customers in partnership with SpringShare LibChat. Last year, the Library delivered over 150 live training sessions via Adobe Connect to over 8,500 students. Working with academic colleagues, the library staff embed library resources and Digital and Information Literacy (DIL) skills materials into modules and qualifications, providing a seamless educational experience (Parker, 2003). 91% of OU students study a module with embedded skills and/or library materials each year. The online delivery model provides the University with large amounts of data, which is explored at The OU in line with the University's Ethical Use of Student Data for Learning Analytics Policy (The Open University, 2014). Research has found that students who use library resources and attend library tutorials get better results (Nurse *et al.*, 2018), (Killick *et al.*, 2018).

## Methodology

Literature in the fields of library performance measurement, library assessment and library data collection practices were reviewed to identify commonalities in practices, focussed primarily on the UK academic library context. The findings were evaluated by theme and have been grouped into four decades of data collection practices:

- 1990s: Performance Indicators
- 2000s: Benchmarking
- 2010s: Value and Impact
- 2020s: Data Intelligence

For each decade the findings outline sector practices and some of their consequences. These are also illustrated by practice case studies from the author.

## Findings

### *100 years of unintended consequences*

Data within libraries is not a new phenomenon. As far back as the 18th century, the Bodleian Library was collecting reader statistics (Killick and Wilson, 2019). In 1919, the University Grants Committee was established by the Government with a remit to allocate funds to universities, including their libraries (Jones, 1986). To support governmental oversight, standardised university statistics were developed covering a wider range of university provision. For libraries, data collected primarily focused on the size of the library collection (e.g. how many books) and the expenditure (e.g. the library budget). When the committee disbanded in 1989 (Shattock, 1994), the Society of College, National and University Libraries (SCONUL) commenced collecting the standardised SCONUL statistics still reported to this day (Creaser, 2009). Although the statistics collected have evolved as our libraries have (Barclay *et al.*, 2012), the size of our collections and our library expenditure is still a key metric we use as a proxy of our quality.

Our consistent datasets tell the story of library evolution. Exploring the past thirty years of book stock statistics we see year-on-year growth in SCONUL libraries from 1993 to 2006 (*table 1*). The number of books in libraries plateaus around 2007 as we enter the global great recession and library budgets come under pressure. By 2013 we start to see a decline in book stock in our collections, as shifts to electronic publishing and pressure on space for study becomes more prevalent. As copyright libraries form part of this dataset with their continued book stock growth, the decline is likely to be more pronounced than the data suggests.

*Table 1: Total catalogued book stock in all UK SCONUL member libraries, 1993 - 2018*

YEAR	TOTAL CATALOGUED BOOK STOCK	DIFFERENCE
93-94	74,406,571	
94-95	77,002,215	3%
95-96	80,155,532	4%
96-97	89,888,557	12%
97-98	93,878,080	4%
98-99	97,282,099	4%
99-00	99,890,543	3%
00-01	101,187,179	1%
01-02	104,967,926	4%
02-03	105,922,015	1%
03-04	108,148,685	2%
04-05	109,649,524	1%
05-06	112,933,304	3%
06-07	113,213,034	0%
07-08	112,970,398	0%
08-09	112,820,493	0%
09-10	111,326,969	-1%
10-11	113,154,811	2%
11-12	111,546,485	-1%
12-13	111,658,764	0%
13-14	109,386,683	-2%
14-15	108,455,142	-1%
15-16	106,568,682	-2%
16-17	102,571,268	-4%
17-18	100,804,223	-2%

What we count has consequences and data has impacts in ways which we cannot always envisage, at times to our own detriment. In the 1920s, faced with an expanding market of academic journals, and their associated costs, libraries started to look for a way to rationalise their subscriptions. Librarians started to look at which journals were most used by readers and ultimately the most cited, prioritising those with higher citations for inclusion in their collections. In turn these developed into the Journal Impact Factors we use today (Gingras, 2016). The academic community is rewarded in terms of career progression (and research funding through the Research Excellence Framework) when published in journals with high impact factors. Publishers in turn are able to charge a premium for titles with high impact factors. Our desire to rationalise our collections in order to reduce our expenditure on journals has resulted in journals with high impact factors dominating academic publishing, and our budgets.

### ***1990s: Performance Indicators***

The performance indicators era of the 1990s focused heavily on input and output measures. A key driver for this within the UK higher education sector was the Jarratt report which recommended to university vice-chancellors that a range of performance indicators should be developed to monitor university efficiency (Johnes and Taylor, 1990). Shortly after the report was published, a series of performance indicators were recommended including 'library stock availability' (Johnes and Taylor, 1990) although this received limited attention from the joint Committee of Vice-Chancellors and Principals (CVCP) / Universities Funding Council (UFC) (Follett, 1993). The Follett report reviewing library provision on behalf of the UK funding councils went on to recommend that a standard series of performance indicators should be developed for university libraries, and that library administrators should use performance indicators as part of their internal management processes (Follett, 1993).

Our focus in data collection moved from input measures (e.g. size of collections and budgets), to efficiency measures. A proxy for efficiency was a series of derived ratios of our input measures represented as output measures. Common examples included stock circulation, calculated by dividing the number of books in a loanable collection by the number of loans in a year. Staff efficiency was derived by dividing the number of enquiries per staff member. The efficiency of the library budget was evaluated by a 'cost per loan' calculation, dividing the cost of the library collection by the number of loans in a period.

Performance indicators were not without consequence. It was possible to pervert the stock circulation statistics by reducing library collections to just stock that had high loan statistics. To improve the metric, book collections were weeded and reduced in favour of popular titles. But what is the purpose of a library? Are you there to guarantee knowledge for the future? Or are you there to ensure that the statistics look good to your paymasters? If we consider that Oakland was defining quality at the same period as meeting the customer's requirements (Beckford, 2009), how did our performance indicators measure the quality of the service we were offering to the person seeking a rare book title in our libraries?

### ***2000s: Benchmarking***

With the development of wide-spread internet use in the 2000s and the ease of data collection and sharing, benchmarking performance became increasingly commonplace. Within libraries, everything started to have an 'e' added, whether that was the content provided through our eJournals, eBooks and eResources, or the methods and tools used to manage and evaluate them via eMetrics and eSurveys.

The E-measures project, led by Conyers and the Evidence Based team, introduced a range of new metrics to the SCONUL Statistics dataset in response to the electronic evolution (Conyers, 2004). The development of COUNTER compliant eJournal statistics provided a trusted, consistent way to see the number of times each packaged journal title had been accessed (Brophy, 2006). The Joint Information Systems Committee (JISC) started to explore the possibility to use usage statistics to inform national negotiations for eJournal packages within the NESLi2 initiative (Bevan *et al.*, 2005).

With easy access to standardised usage statistics, we started to benchmark our eJournal collections against each other within our individual library contexts. Our output measures evolved from cost per loan to cost per download. Journal titles were

evaluated against one another to identify which provided the best ‘value for money’ based on cost per download, and we agonised over the ‘long tail’ of journal titles in the deal with low usage (Killick, 2013).

Our measurement behaviour here also had consequences. Cost per download analysis naturally favoured journals with a large, broad readership and disciplines where the university had higher populations studying that area. Titles offered by smaller publishers in niche fields with fewer students within the university were viewed as being low-demand, attracting a higher cost per download. Again, this causes us to question the purpose of the library. If a journal title costs £5,000 per year, and receives one article access in the year, is it a waste of money for the library? What if that one access enabled a £15m research project to succeed? What if that one access was crucial in informing the development of a Covid vaccine? Is the £5,000 a useful investment then?

Alongside evaluating our collections, we also sought to benchmark our customer satisfaction. The ease of distribution and analysis of online surveys, and the desire to move away from efficiency measurement saw the development of LibQUAL+ (Cook and Heath, 2001). The team at Texas A&M University worked with the Association of Research Libraries to develop the standardised library quality survey, which has been used by thousands of libraries worldwide – including over 70 SCONUL members over the years (Killick and Town, 2012). The survey provides respondents with the option to rate the library in terms of minimum and desired service levels, alongside perceived performance for 22 standard questions. Questions focus on three dimensions of library service – the support provided by staff, the library collection and ease of access to information, and the physical library space. With the same questions being used by all participants, libraries had the opportunity to evaluate their performance against that of their peers and sector benchmarks.

In 2005, the National Student Survey (NSS) was introduced, providing the first mandated student satisfaction survey for UK Higher Education aimed at final year undergraduates (Richardson *et al.*, 2007). Covering all aspects of the education experience for students, one question was presented covering the library offer: ‘the library services and resources are good enough for my needs’ (Killick and Wilson, 2019). Unlike LibQUAL+, whose data was only visible to library administrators, the NSS results provided the public with a league table of university performance (Richardson *et al.*, 2007). The visibility of the results led to universities viewing the NSS results as business critical. For library administrators, league tables seemed to benefit those at the top for promotional reasons, and those at the bottom who were able to leverage the results to unlock funding (Stanley, 2009).

But are we looking to prove, or improve? Standardised surveys do not account for different forms of service provision tailored to local objectives (Creaser, 2006). The Open University’s embedded library strategy, where content and Digital and Information Literacy (DIL) skills were seamlessly woven into learning materials, meant that the Library was an invisible partner to the students. This approach is pedagogically sound and it met student learning needs; why should the student care that it is the Library providing their skills materials and online reading list resources? But when the NSS revealed that 20% of OU students perceived that they did not have a library, by answering ‘not applicable’ to the library NSS question, the Library had to change its strategy. The visibility of the Library in the students’ learning became increasingly important. Learning materials were rewritten to emphasise that skills materials and resources were being created and delivered by the Library. Online information architecture projects were created to improve Library visibility. Library marketing strategies were developed to inform students of its vital role in their education. As a result of this strategy, the number of students answering ‘not applicable’ to the library NSS question has reduced to 2%, and satisfaction scores have increased. These actions have been done to prove the value of the Library, not necessarily to improve the student experience.

### ***2010s: Value and Impact***

In the Value and Impact Era of the 2010s, we started to see data break out of silos, linking student attainment with library datasets. Several global projects looking at correlation data emerged, with the University of Huddersfield pioneering this work in the UK (White and Stone, 2010). Positive correlations between library use and student success were found, and we started to explore if there was a statistically significant correlation between library activity data and student attainment in university libraries (Stone *et al.*, 2011).

Technology had enabled us to link a variety of different activity datasets with student attainment results. Over the past 10 years, we have seen a variety of studies seeking to identify the relationship between students' access to electronic resources, borrowing of print books, visits to the library building, and attendance at training sessions; all linked to student attainment (Killick *et al.*, 2018; Nurse *et al.*, 2018; Soria *et al.*, 2013).

One of the drivers for doing this is to evidence to senior stakeholders the impact the library has on the university endeavour. Within any context, the library is always responsible to its parent organisation and competing for funding. You may have a Vice Chancellor who intrinsically knows the complexities of libraries and the value they offer the organisation, without the need for any evidence to be shown to prove this. Chances are, however, in a world where resources are scarce and are needed by many, your ability to show evidence of impact is going to be highly valuable.

Again, this is not without consequence. Issues of data ethics and privacy have been challenging librarians during the past decade (Asher, 2017). Within library learning analytics, large datasets where individuals are not identifiable have been used to evaluate the impact of the library on students in line with data ethics policies (Killick *et al.*, 2018; Nurse *et al.*, 2018). With aspirations for improving equality, diversity and inclusivity in higher education, using mass datasets to inform service improvements negates the needs of the individual, and we run the risk of continuing to exclude rather than include. The datasets are primarily used to prove, rather than to improve.

### **2020s: Data Intelligence**

Looking ahead, as we enter an era of data intelligence, we have more data available to us than ever before which we use to improve our services.

At The Open University, the Library uses a Customer Relationship Management (CRM) system to manage all enquiries received at the helpdesk. These are categorised by the Library staff by on the nature of the enquiry, combined with the customer record and the contact route (e.g. webchat, email, telephone, in-person). The Library team can analyse the enquiries data to design targeted support for the courses which contact the Library helpdesk the most. They can also identify the most frequent areas of enquiry across all courses, informing proactive support materials and training sessions. It can also highlight where the Library needs to improve its services to meet the needs of individual students.

The Library also uses the CRM data to forecast workload for the year ahead. Based on average handling times for enquiries by each contact route, forecasts are created detailing the amount of time needed by the team to answer the anticipated enquiries. Forecasts are designed for every day of the year, as the number of enquiries fluctuate by the day of the week and the week in the year. The helpdesk does have a minimum staffing level which is always maintained, however, through forecasting, the team are able to predict when to increase the staffing level across the year. This enables the Library to ensure staffing levels are right, providing responses to students which meet the service level agreement, and freeing up staff for other work at other times of the year.

The OU Library has more data on its customers than ever before. It has more opportunities to use that data to improve the services it provides; and to prove its value. There are still many unanswered questions data could inform. Do Library group live training sessions result in a reduction in Library enquiries? Are Library embedded skills materials having a positive impact on student success? Do courses with larger amounts of Library content and skills embedded in them have greater student satisfaction? Does providing Library content and embedded skills development reduce attainment gaps?

[But] management through data alone is dangerous. If we are defining quality as meeting or exceeding customer satisfaction with the product or service, data alone is no indicator of quality. Deming (2000) in his Seven Deadly Diseases of Management discussed the risk of 'management by use only of visible figures, with little or no consideration of figures that are unknown or unknowable'. He recognised the intangible aspects of services, and that what can be measured is only a 'trivial part of the gain' of an organisation (Beckford, 2009).

Which is where our professional expertise comes in. This is something we all have in our libraries but rarely discuss. Our biggest asset, and most valuable resource, is the library team. We enable and empower our users to become information literate. We collaborate and innovate our services, continually exploring more possibilities to preserve and unlock

knowledge. We offer our universities unique professional expertise about knowledge and information management, digital capabilities and asset preservation and reuse. Generally, library professionals like to quietly get on with the job in an unassuming manner, but if we do not shout about our expertise, value, and impact, who will? Using data to prove our value, and communicating it as a loud librarian (Killick and Wilson, 2019) is essential to our success and survival; providing we are mindful of the consequences of our actions.

## Conclusion

Libraries have always used data and will always continue to do so. It is a powerful tool in our quest to both prove our value and improve our services. Our data practices have evolved in line with societal thinking and sector practice, from counting the size of library collections to the library learning analytics of today. What we measure matters and the consequences of our use of the data cannot always be predicted. Ensuring we continue to use data in a safe, ethical manner to support our service objectives is vital. We must also acknowledge that not everything can be counted: some key strengths and value offered by our libraries will be unknown, and will remain unknowable.

## References

- Asher, A.D. (2017), "Risk, Benefits, and User Privacy: Evaluating the Ethics of Library Data.", in Newman, B. and Tijerina, B. (Eds.), *Protecting Patron Privacy: A LITA Guide.*, Rowman & Littlefield, Lanham, MD, pp. 43–56.
- Barclay, P., Conyers, A., Creaser, C. and White, S. (2012), "Performance measurement in a changing environment", *Performance Measurement and Metrics*, Vol. 13 No. 2, pp. 92–106.
- Beckford, J. (2009), *Quality: A Critical Introduction*, Third., Taylor & Francis Group, London.
- Bevan, S., Dalton, P. and Conyers, A. (2005), "How usage statistics can inform national negotiations and strategies", *Serials*, Vol. 18 No. 2, pp. 116–123.
- Brophy, P. (2006), *Measuring Library Performance: Principles and Techniques*, 1st ed., Facet Publishing, London.
- Conyers, A. (2004), "E-measures: developing statistical measures for electronic information services", *Vine*, Vol. 34 No. 4, pp. 148–153.
- Cook, C. and Heath, F. (2001), "Users' Perceptions of Library Service Quality: A LibQUAL+ Qualitative Study", *Library Trends*, Vol. 49 No. 4, pp. 548–584.
- Creaser, C. (2006), "One size does not fit all: user surveys in academic libraries", *Performance Measurement and Metrics*, Vol. 7 No. 3, pp. 153–162.
- Creaser, C. (2009), "UK Higher Education Library Statistics", in Heaney, M. (Ed.), *Library Statistics for the Twenty-First Century World: Proceedings of the Conference Held in Montréal on 18-19 August 2008 Reporting on the Global Library Statistics Project*, Berlin, New York: K. G. Saur, 2009. <https://doi.org/10.1515/9783598441677>.
- Deming, W.E. (2000), *Out of the Crisis*, First MIT Press Edition, MIT Press, Cambridge, Massachusetts.
- Follett, B. (1993), *Joint Funding Councils' Libraries Review Group: Report*, Higher Education Funding Council for England, Bristol.
- Gingras, Y. (2016), "The Origins of Bibliometrics", *Bibliometrics and Research Evaluation: Uses and Abuses*, MIT Press, pp. 1–10.
- Johnes, J. and Taylor, J. (1990), *Performance Indicators in Higher Education.*, Society for Research into Higher Education & Open University Press, Buckingham.
- Jones, G. (1986), "But how fast did topsy grow?", *Library Review*, Vol. 35 No. 2, pp. 110–114.
- Killick, S. (2013), "Applying Performance Measurement to Safeguard Budgets: Qualitative and Quantitative Measurement of Electronic Journal Packages", in Hiller, S., Kyrillidou, M., Pappalardo, A., Self, J. and Yeager, A. (Eds.), *Proceedings of the 2012 Library Assessment Conference. Building Effective, Sustainable, Practical Assessment*, Association of Research Libraries, Charlottesville, pp. 731–735. Available at: <http://oro.open.ac.uk/45855/>
- Killick, S., Nurse, R. and Clough, H. (2018), "The Continuing Adventures of Library Learning Analytics: Exploring the Relationship between Library Skills Training and Student Success.", in Baughman, S., Hiller, S., Monroe, K. and Pappalardo, A. (Eds.), *Proceedings of the 2018 Library Assessment Conference. Building Effective, Sustainable, Practical Assessment.*, Association of Research Libraries, Houston, TX, pp. 188–199. Available at: <http://oro.open.ac.uk/66217/>
- Killick, S. and Town, J.S. (2012), "LibQUAL +: The SCONUL experience", *Sconul Focus*, Vol. 54, pp. 29–32. Available at: <http://oro.open.ac.uk/45856/>
- Killick, S. and Wilson, F. (2019), *Putting Library Assessment Data to Work*, Facet Publishing, London.

- Nurse, R., Baker, K. and Gambles, A. (2018), "Library resources, student success and the distance-learning university", *Information and Learning Science*, Vol. 119 No. 1–2, pp. 77–86. Available at: <http://oro.open.ac.uk/49730/>
- Parker, J. (2003), "Putting the pieces together: information literacy at The Open University", *Library Management*, Vol. 24 No. 4/5, pp. 223–228. Available at: <http://oro.open.ac.uk/44039/>
- Richardson, J.T.E., Slater, J.B. and Wilson, J. (2007), "The National Student Survey: Development, findings and implications", *Studies in Higher Education*, Vol. 32 No. 5, pp. 557–580.
- Shattock, M. (1994), *The UGC and the Management of British Universities*, Society for Research into Higher Education & Open University Press, Buckingham.
- Soria, M.K., Fransen, J. and Nackerud, S. (2013), "Library Use and Undergraduate Student Outcomes: New Evidence for Students' Retention and Academic Success", *Portal: Libraries and the Academy*, Vol. 13, pp. 147–164.
- Stanley, T. (2009), "The National Student Survey: Pain or Gain?", *SCONUL Focus*, Vol. 45, pp. 144–147.
- Stone, G., Ramsden, B. and Pattern, D. (2011), "Looking for the link between library usage and student attainment", *Ariadne*, No. 67, available at: <http://eprints.hud.ac.uk/10774/>.
- The Open University. (2014), *Policy on Ethical Use of Student Data for Learning Analytics*, Milton Keynes, available at: <http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ethical-use-of-student-data-policy.pdf>.
- White, S. and Stone, G. (2010), "Maximizing use of library resources at the University of Huddersfield", *Serials*, Vol. 23 No. 2, pp. 83–90.

#### **Author Biography:**

Selena Killick presents, publishes, and provides consultancy services to libraries on an international basis on using assessment methods to inform library strategy and improve service delivery. Her research focus has been on library performance measurement and assessment since 2003. She is the co-author of *Putting Library Assessment Data to Work*, and is a member of the Library Performance Measurement Conference Board. She is the Associate Director within Library Services at The Open University, UK. Her remit includes leading the development and delivery of strategies for the Library to meet the needs of the OU community.

You can contact her via <https://selenakillick.com/>